Education in OECD's PIAAC Study: How Well do Different Harmonized Measures Predict Skills?

Schneider, Silke L.

methods, data, analyses | Vol. 12(1), 2018, pp. 151-176 DOI: 10.12758/mda.2017.15

Education in OECD's PIAAC Study: How Well do Different Harmonized Measures Predict Skills?

Silke L. Schneider
GESIS – Leibniz Institute for the Social Sciences

Abstract

The comparable measurement of educational attainment is a challenge for all comparative surveys and cross-national data analyses. While education is an important predictor or control variable in many research contexts, it is particularly important when studying education and education-related outcomes such as skills or labor market chances. This study evaluates the cross-nationally comparable measurement of education in OECD's Programme for the International Assessment of Adult Competencies, PIAAC, in terms of its construct validity when predicting general basic skills. In order to do so, the predictive power of country-specific (i.e. non-comparable) education variables is compared to the predictive power of different cross-nationally harmonized variables, namely the detailed ISCED-based coding scheme used in PIAAC, ISCED 2011 and 1997 levels, the broad education levels 'low, medium, high', ES-ISCED, as well as years of education. The analyses consist in sets of country-wise linear regressions, taking PIAAC's plausible values and complex sampling into account, and use adjusted R^2 as the indicator for predictive power and validity. The results show that while harmonization into a detailed coding scheme such as the most detailed comparable variable available in PIAAC does not entail large losses of information, the way this variable is further simplified plays a major role for validity. The paper also highlights shortcomings of the detailed variable from a theoretical point of view, such as the lack of differentiation of vocational and general education and other markers of educational content and quality, which are important aspects both for skill development as well as the labor market outcomes of education, and of the country-specific measures of education, which may make the detailed PIAAC education variable look better than it actually is.

Keywords: Measurement; Educational attainment; Skills; Comparative research; Education; Data quality; Survey

Introduction

An important challenge in comparative survey research is how to make data comparable or 'functionally equivalent' across countries (Przeworski & Teune 1970). The underlying process is called 'harmonization' (Wolf et al., 2016; Hoffmeyer-Zlotnik & Wolf, 2003), especially when speaking about the comparability of individual variables (rather than e.g. sampling or fieldwork procedures). Harmonizing survey data cross-nationally entails the risk of 'harmonizing away' meaningful information (Granda et al., 2010). When a harmonized variable carries less information than a non-harmonized one, and the amount of information loss differs across countries, the comparability of the harmonized measure is necessarily limited. This is an important element of comparison error (Smith 2011), a main impediment of successful comparative survey research.

The comparability of background variables such as ethnicity, education or social class (see e.g. Schneider et al., 2016; Braun & Mohler, 2003) has mostly been researched regarding the education variable. This is for two reasons: Firstly, education is a major independent variable in numerous statistical models of survey micro data, either as control or substantive variable, and thus maybe the most important of all background variables (Smith, 1995). Secondly, its harmonization is, because of the stark institutional differences between educational systems, particularly difficult (Braun & Müller, 1997). Cross-national educational attainment levels such as ‚primary education' or ‚tertiary education', even if translated correctly, are likely to be interpreted differently by respondents in different countries depending on features of their educational systems. Therefore, the state of the art for cross-national surveys is to use country-specific questionnaire items to collect information on respondents' educational attainment (Schneider, 2016). The resulting country-specific education variables are then recoded into a cross-national variable after data collection. This approach is called ex-ante output harmonization (Wolf et al., 2016; Ehling, 2003). Today, most surveys use UNESCO's International Standard Classification of Education (ISCED, UNESCO Institute for Statistics, 2012) for harmonizing education variables.

However, there is no agreement on which specific ISCED-based variables to provide to data users – three broad levels, main ISCED levels, or whether sub-categories within levels representing different types of education also need to be taken into account. The method of comparative construct validation is fairly established today for evaluating the comparative validity of harmonized education variables

Direct correspondence to
Silke L. Schneider, GESIS – Leibniz Institute for the Social Sciences,
P.O. Box 122155, 68072 Mannheim, Germany
E-mail: silke.schneider@gesis.org

in cross-national survey data. These analyses consist in sets of country-wise linear regressions, and usually use adjusted R^2 as the indicator for predictive power or validity. Prior research using this method (Schneider, 2010; Kerckhoff & Dylan, 1999; Kerckhoff et al., 2002; Kieffer, 2010; Müller & Klein, 2008; Braun & Müller, 1997) has generally concluded that the education variables in comparative surveys, including those based on ISCED, contain comparison error, especially (but not exclusively) due to the way that country-specific education categories are aggregated into supposedly comparable, broader categories.

This paper adds to this research using the OECD's Programme for the International Assessment of Adult Competencies, PIAAC (OECD, 2013; OECD, 2016a). In addition to not having been the object of a comparative construct validation of the harmonized education variable yet, PIAAC also offers new validation variables that have so far not been exploited for a comparative construct validation, namely literacy and numeracy skills. The relationship between educational attainment and skills is expected to be fairly strong (and thus sensitive to measurement quality) because one important aim of formal education and training systems is skill production (see e.g. Hall & Soskice, 2001). Because of the close relationship between educational attainment and skills, if educational attainment is not well measured, in statistical models using both as independent variables, unmeasured heterogeneity in education may be picked up by the measure of skills (confounding). It is thus of great importance in a survey of adult skills that educational attainment is measured with a high degree of reliability and validity. Such an analysis will also help us better understand the relationship between educational qualifications and skills (Heisig & Solga, 2015).

This paper builds on the work by Schneider (2010), which used occupational status as the validation variable, and evaluates the harmonized educational attainment measures employed in PIAAC. It answers the following research questions:

1. How comparable across countries, in terms of comparative validity, is the most detailed comparative education variable provided in the PIAAC data set?
2. Do we find the same result for the PIAAC data that were previously found for the ESS and other surveys, namely that main education levels and nominal years of education diminish comparative validity? How does ISCED 2011 fare, compared with ISCED 1997?
3. Could a differently aggregated education variable, such as the European Survey Version of ISCED (ES-ISCED) proposed in Schneider (2010), improve the comparative validity of education measures in PIAAC?

This paper starts out by distinguishing dimensions of education and theorizing about their relationship with general basic skills. Then, the PIAAC data and analysis methods will be presented, as well as the harmonized measures of education available in PIAAC. Here the implications of the theoretical rationale for the measurement of educational attainment are also presented. After presenting the empirical results, the paper will summarize and conclude with some practical recommendations for the next Cycle of PIAAC, which will also be relevant to other future cross-national surveys as well as research using existing data.

Dimensions of Education and General Basic Skills

From a theoretical point of view, education and skills are expected to be fairly closely related. In modern societies, formal education is an important source of general basic skill development and 'human capital' (Becker, 1964; OECD, 2013; OECD, 2016a). Examinations in formal education aim to validate the successful acquisition of knowledge, skills and competences, and give legitimacy to subsequently achieved advantageous social positions (Weber, 1922). Consequently, formal educational qualifications are the most common indicator for educational attainment in surveys.

Formal education is not homogeneous but differs in terms of quality, content and type in very complex ways (Smith, 1995). The educational systems in most developed countries provide alternative programs within education levels. Depending on their specific goals and curricula, different types of educational programs can be expected to lead to different levels of general basic skills. In the following, the dimensions of education distinguished by Smith (1995) - quantity, content, quality and type - are examined with respect to their implications for general basic skills, and hypotheses formed for measuring education in such a way as to optimize the prediction of skills by education.[1]

The first dimension of education is *quantity*. From a human capital point of view (Becker, 1964), the longer children go to school, and the higher the level of education eventually reached by youth and young adults, the stronger we expect their literacy and numeracy skills to be. The better an education measure reflects quantity, the better it is thus expected to predict general basic skills (Hypothesis 1).

The second dimension is *content*, i.e. "what is being taught" (Smith, 1995, p.218). Some (especially European) countries track children in lower secondary education already into programs with different content, preparing for different labor market 'careers' (Haller et al., 1985; Braun & Müller, 1997; König et al., 1988). From upper secondary education onwards, *most* (if not all) countries offer different educational programs with specialized content, mostly differentiating university preparatory general education and vocational programs preparing for the labor market. In vocational education, students spend some of their time learning

1 It is important to note that we simplify these dimensions substantially here, compared to the rich array of indicators that Smith himself has to offer for each of them.

practical skills directly relevant to the labor market. In contrast, in general education, most learning time is spent on text and number based tasks. Therefore, at the simplest level, the better an education measure distinguishes between vocational and general programs, the better it is expected to predict general basic skills (Hypothesis 2).

The third dimension of education is *quality*. In countries offering different educational programs or institutional settings at any single level of education, these may differ not only with respect to their curricular content but also their skill (and social) selectivity, an important indicator for the quality of education (Smith, 1995). For example, many countries especially in Eastern Europe have different types of vocational upper secondary education programs (see e.g. Saar, 2008; Straková, 2008; Bukodi et al., 2008). Some of them give access to higher education, while others do not. Typically, those providing access to higher education are more selective and academically demanding, while those only preparing for the labor market are less so. This results in higher skills of graduates from the former programs, which are however typically already evident when *entering* the program and are thus not or only partially (e.g. through the dimension of content, see above) *caused* by the program. A similar argument can be made for tracking in lower secondary education, where different programs may work at different standards. We thus expect education measures that differentiate educational categories by skill selectivity or institutional setting to better predict adult skills than measures not making such a distinction (Hypothesis 3).[2]

The fourth dimension of education according to Smith (1995) is *type*, which consists in several distinct classification systems that partly overlap with those previously discussed. A distinction by type not yet covered but useful here is the place of learning, where we can distinguish entirely school-based programs from programs combining schooling and work, as in apprenticeship programs in mostly German-speaking countries, and on the job training (see e.g. Allmendinger, 1989), where the latter does not count as formal education. Because of the more strongly theoretical content and book-based learning, we can expect the completion of school-based vocational programs to be related to higher general basic skills than apprenticeship programs, where practical learning plays a more prominent role. Therefore, education measures distinguishing between school-based and apprenticeship programs are expected to better predict general basic skills than measures not making such a distinction (Hypothesis 4).

2 In many countries, content and quality of education are overlapping dimensions: academically or generally oriented programs are usually more selective and provided in specialized institutional settings (such as the prototypical Gymnasium or traditional university), while vocationally or professionally oriented programs are - at least at the secondary level - less selective and, in countries with differentiated vocational training systems, provided in a variety of institutional settings.

To summarize, a valid comparable measure of educational attainment that well reflects skills may need to differentiate types of formal education in addition to levels of education, ideally in terms of tracks in lower secondary schooling, programme orientation, and, especially within vocational education, selectivity and place of learning. Measures that simplify education by reducing it to one dimension, such as broad levels of education or duration in terms of years of education, can be expected to function less well and less consistently across countries in predicting general basic skills. The dimensions discussed here may also help explain *why* country-specific measures sometimes do a better job at predicting general basic skills than comparative measures.

Data and Methods

The Programme for the International Assessment of Adult Competencies

OECD's Programme for the International Assessment of Adult Competencies (PIAAC) is a cross-national large-scale survey assessing the general basic skills of the adult population – literacy, numeracy and problem solving in technology-rich environments – that are considered essential for successful participation in today's societies (OECD, 2016a). While skills are directly assessed using psychological tests, information on demographic characteristics, education, labor market participation and other indicators are collected using a background questionnaire. Data for the first set of countries (round 1, 24 countries[3]) were collected in 2011/2012, and for a second set of countries (round 2, 9 countries[4]) in 2014/2015. The target population consisted of individuals aged 16 to 65. Multi-stage random sampling techniques with complex sampling designs were employed. Samples sizes range from just below 5000 (minimum requirement) to about 21000 (Canada). Further details are available in the technical report (OECD, 2016b).

For the analyses in this paper, individuals under age 30 are only included if they are not currently in formal education. Respondents who obtained their highest educational qualification abroad are excluded because a high degree of measure-

3 Australia, Austria, Belgium (Flanders only), Canada, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Ireland, Italy, Japan, Korea, the Netherlands, Norway, Poland, Russia (excluding the Moscow municipal area), the Slovak Republic, Spain, Sweden, the United Kingdom (England and Northern Ireland only) and the United States.

4 Chile, Greece, Indonesia (Jakarta only) Israel, Lithuania, Singapore, New Zealand, Slovenia and Turkey. Data for Indonesia have not been released. For Greece, about a fifth of cases did not have responses for the direct assessments. These were imputed by OECD.

ment error on the educational attainment variable can be expected for these respondents.

Education Variables to be Compared Across Countries

Looking at survey practice, different cross-national surveys and analyses use different coding schemes, even if they refer to ISCED. ISCED primarily distinguishes levels of education, ranging from less than primary education to the PhD level. In order to distinguish between attainment of different *types* of education, ISCED allows education to be differentiated, within levels, by programme orientation (general vs. vocational) and whether a qualification gives access to a higher education level or not. The details of these distinctions have somewhat changed between ISCED 1997 to 2011 (see Schneider, 2013; UNESCO Institute for Statistics, 2012). In PIAAC, a coding scheme closely related to the implementation of ISCED 97 in the European Union Labor Force Survey (EU-LFS) until 2013 was used (variable name *B_Q01a*, see first column of Table 1). This coding scheme differentiates educational programs at the upper secondary level not allowing access to tertiary education (ISCED 3C, usually vocationally oriented) from programs giving such access (ISCED 3A-B, which may be generally or vocationally oriented). In PIAAC, the Bachelor and Master levels are additionally distinguished from short vocational tertiary education, anticipating ISCED 2011. Compared to other surveys, this is a fairly detailed coding scheme. The following less detailed ISCED-based variables are also included in the validation:

- *ISCED 2011 levels*, derived from *B_Q01a* (9 categories).
- *ISCED 1997 levels*, also derived from *B_Q01a* (7 categories).
- *Broad ISCED levels* represent a further aggregation, resulting in three education levels: less than upper secondary (low), upper secondary including post-secondary non-tertiary (medium), and tertiary (high). This coding is commonly used in statistical reporting and cross tabulations, but also in multivariate analyses.

Table 1 shows how these different variables relate to each other.

Table 1 ISCED coding schemes available in PIAAC data or derived

<table>
<tr><th>B_Q01a</th><th>ISCED 97</th><th>ISCED 11</th><th>Broad ISCED</th></tr>
<tr><td>0 No formal qualification or below ISCED 1</td><td colspan="2">0 No formal qualification or below ISCED 1</td><td rowspan="3">1 low</td></tr>
<tr><td>1 ISCED 1 (primary education)</td><td colspan="2">1 ISCED 1 (primary education)</td></tr>
<tr><td>2 ISCED 2
3 ISCED 3C <2 years</td><td colspan="2">2 ISCED 2 (lower secondary)</td></tr>
<tr><td>4 ISCED 3C 2 years+
5 ISCED 3A-B
6 ISCED 3 (no distinction A-B-C)</td><td colspan="2">3 ISCED 3 (upper secondary)</td><td rowspan="2">2 medium</td></tr>
<tr><td>7 ISCED 4C
8 ISCED 4A-B
9 ISCED 4 (no distinction A-B-C)</td><td colspan="2">4 ISCED 4 (post-secondary non-tertiary)</td></tr>
<tr><td>10 ISCED 5B
11 ISCED 5A, bachelor level
12 ISCED 5A, master level</td><td>5 ISCED 5 (tertiary 1)</td><td>5 ISCED 5
6 ISCED 6
7 ISCED 7</td><td rowspan="2">3 high</td></tr>
<tr><td>13 ISCED 6 (tertiary 2)</td><td>6 ISCED 6 (tertiary 2)</td><td>8 ISCED 8</td></tr>
</table>

An alternative and very popular indicator of educational attainment is *years of education*, a generalization of the 'years of schooling' prominently used by Blau and Duncan (1967). In contrast to the other comparative measures, this is a linear variable. In this study, hypothetical years of education are derived from national measures of the highest educational qualification obtained by assigning nominally required years of education to educational qualifications. In PIAAC, such a variable is provided (variable name *yrsqual*).

This study also evaluates the European Survey version of ISCED (ES-ISCED) proposed in Schneider (2010), which was developed in order to integrate some basic ideas underlying CASMIN[5] in data coded with ISCED. This variable aims to minimize loss of information through harmonization by including a minimal degree of

5 The CASMIN education scheme (König et al., 1988) is used a lot for ex-post harmonization of country-specific education variables in surveys (see e.g. Breen et al., 2009; Müller & Karle, 1993). CASMIN cannot be coded for PIAAC because we lack respective documentation for a large number of PIAAC countries, and for many countries, the country-specific variables are not differentiated enough to allow coding into CASMIN.

within-levels differentiation in terms of educational content and quality, while not being more detailed than ISCED 97 main levels, by aggregating main levels that are typically very small in European (and likely most developed) countries. Table 2 shows how it was derived, for the purpose of this study, from *B_Q01a* and the additional indicator variable *VET* (for vocational education and training).[6] Some distinctions that would have been necessary for the construction of ES-ISCED could not be made in PIAAC, so that ES-ISCED here only approximates ES-ISCED as proposed in Schneider (2010).

While none of these comparative education measures covers all dimensions presented in section 2, and such a measure also could not be constructed from PIAAC data, we can still form some expectations based on the above hypotheses. *B_Q01a* reflects skill selectivity to some degree at both secondary and tertiary levels using destination (A, B and C), but does not explicitly reflect orientation, which to some degree however overlaps with destination. The aggregated ISCED variables reflect the duration of education in a more or less differentiated way, but neither program orientation nor skill selectivity. Years of education focus on quantity exclusively. ES-ISCED most strongly reflects the distinction between general and vocational content but sacrifices quantity at the lowest and highest levels. Following hypothesis 1 (quantity), we thus expect the following order of the measures in terms of performance predicting skills: years of education > B_Q01a > ISCED 2011 levels > ISCED 1997 levels > ES-ISCED > broad ISCED levels. Regarding hypothesis 2 (content: vocational vs. general orientation), we expect ES-ISCED to perform better than all other measures except maybe B_Q01a. Hypothesis 3 (quality: institutional and selectivity differentiation) makes us expect B_Q01a to perform best, especially as regards the distinction within vocational programs in Eastern European countries in ISCED 3C vs. ISCED 3A-B, followed by ES-ISCED. Hypothesis 4 (type: school-based vs. apprenticeship) is not operationalized in either comparative variable but visible in some country-specific variables.

6 *VET* was coded centrally in PIAAC *after* data collection and aims to provide a differentiation between general and vocational education at ISCED levels 3 and 4. Unfortunately, the variable *VET* contains a large amount of missing data even for countries where the educational system visibly distinguishes between vocational and general education. These countries did not distinguish vocational and general education in their educational attainment measures because they were not required to do so when the country-specific education measures for PIAAC were designed. For these, it was thus impossible to provide this information ex-post. Therefore, firstly a close examination of country-specific variables and the *VET* variable was conducted so as to correct some codings in *VET*, and secondly a new category IIIu for remaining unspecified orientation at this level was added to ES-ISCED.

Table 2 Correspondence between PIAAC variables B_Q01a, VET, and ES-ISCED

B_Q01a	Label	VET	ES-ISCED
1	No formal qualification or below ISCED 1	-	I
2	ISCED 1	-	
3	ISCED 2	-	II
4	ISCED 3C shorter than 2 years	-	
5	ISCED 3C 2 years or more	0 (general) or missing	
5	ISCED 3C 2 years or more	1 (vocational)	IIIb[1]
6	ISCED 3A-B		
7	ISCED 3 (without distinction A-B-C, 2y+)		
6	ISCED 3A-B	0 (general)	IIIa[2]
7	ISCED 3 (without distinction A-B-C, 2y+		
9	ISCED 4A-B		
10	ISCED 4 (without distinction A-B-C)		
6	ISCED 3A-B	missing	IIIu
7	ISCED 3 (without distinction A-B-C, 2y+		
9	ISCED 4A-B		
10	ISCED 4 (without distinction A-B-C)		
8	ISCED 4C	any	IV[3]
9	ISCED 4A-B	1 (vocational)	
10	ISCED 4 (without distinction A-B-C)		
11	ISCED 5B	-	
12	ISCED 5A, bachelor degree	-	V1
13	ISCED 5A, master degree	-	V2
14	ISCED 6	-	

Notes.
1 This category should have included ISCED 3B general but not ISCED 3A vocational, which however cannot be identified in PIAAC.
2 This category should have included ISCED 3A vocational but not ISCED 3B general, which however cannot be identified in PIAAC.
3 This category should have included 4B general, which however cannot be identified in PIAAC

Comparative Construct Validation Method

In order to evaluate the loss of information and validity caused by the harmonization of country-specific education variables into various comparative education variables across countries, and thus to find out which kind of comparative education coding scheme best represents the information contained in country-specific measures in terms of basic skills, PIAAC data are subjected to a series of linear regression analyses by country, following Schneider (2010). Literacy skills are used as the validation (dependent) variable here, but the results look very similar when using numeracy rather than literacy skills as validation construct (see Figure 4 and Table 9 in the appendix). The first or benchmark model uses the country-specific education variables, coded as dummies, as the main predictor.[7] The subsequent models use the comparative education variables described above, also coded as dummies. Years of education are treated as a linear variable. All models control for sex and age.

The measure of predictive power or information preserved in the harmonized variable is the relative adjusted R^2 of the respective model in comparison with the benchmark model, i.e. the adjusted R^2 of the model using the comparative education variable to be evaluated as predictor divided by the adjusted R^2 of the benchmark model using the country-specific education variable as predictor. This relative view on losses of information takes into account that the overall association between education and skills differs across countries, and that the same absolute reduction in predictive power is more severe at lower levels of association than at higher levels. Absolute losses in R^2 are reported in the appendix (Table 8). The R^2s are multiplied by 100 to allow a percentage interpretation. In all models, both the complex survey design in PIAAC as well as the representation of skills as 'plausible values' are taken into account. The analyses were performed in Stata 14 using the Stata package 'repest' (Avvisati & Keslair, 2017).

To further facilitate interpretation, cross-country statistics are calculated. In order to check whether individual comparative education variables lead to higher or lower variation in predictive power across countries, standard deviations are also reported. High variation in relative predictive power across countries means that a harmonized variable does not work equally well across countries, thus threatening comparability.

7 These are not available in the public use files and thus required analyzing the data at OECD. Australia did not provide country-specific source variables to OECD. Therefore, *B_Q01a* is used as the benchmark for Australia, so that for this country, only the performance of comparative variables relative to the most detailed comparative variable can be evaluated. Some countries used several questionnaire items for measuring educational attainment. These were combined into one country-specific variable before analysis.

Results

The results of the analyses are presented in three steps: Firstly, before interpreting the results of the comparative education variables, it is worth looking at the results concerning the country specific variables. If these do not highly correlate with literacy skills as expected by theory, one may be skeptical with regards to their measurement quality, putting their usefulness as a quality benchmark into doubt.[8] Secondly, to get an idea of how different harmonized education variables work, we look at the summary statistics regarding the relative predictive power of these variables compared to the country-specific variables. Thirdly, the paper takes a more detailed look at the regression coefficients in the benchmark model for selected countries where the biggest problems were identified in the previous step. This is the strategy also followed by Müller and Klein (2008) for Germany in EU-SILC.

The Benchmark Model

The R^2s representing the strength of association between country-specific education measures and skills, including effects of sex and age, resulting from the benchmark model are shown in Figure 1. Some countries show unexpectedly weak relationships even when using country specific education variables. These are Russia (4% adjusted R^2), Cyprus and Greece (each 12%), Lithuania (16%) and Estonia (19%). While the results for the Baltic states may not be entirely off, we should be careful interpreting the results for these countries: either the country-specific measurement instruments are of low quality already, or there are other data quality issues involved. Other countries in contrast show strong links between educational attainment, sex, age and skills, which is closer to what is theoretically expected. In Singapore, sex, age and education explain more than 50% of the variation in literacy skills, followed by the Netherlands with 40%. Flanders, Chile, France and French-speaking Canada all have 36-37%. Beyond having better education measures, the effects of sex and especially age may also be stronger in these countries.

Validity of Comparative Education Variables

The results of the analysis comparing the performance of comparative education variables with country specific ones are shown in Figure 2 (selected summary statistics) and Table 3 (detailed results for all countries and summary statistics). Adjusted R^2s are shown relative to those reported in Figure 1, which are thus set

8 Of course, some degree of ‘real’ cross-national variation in the relationship between education and skills is also to be expected.

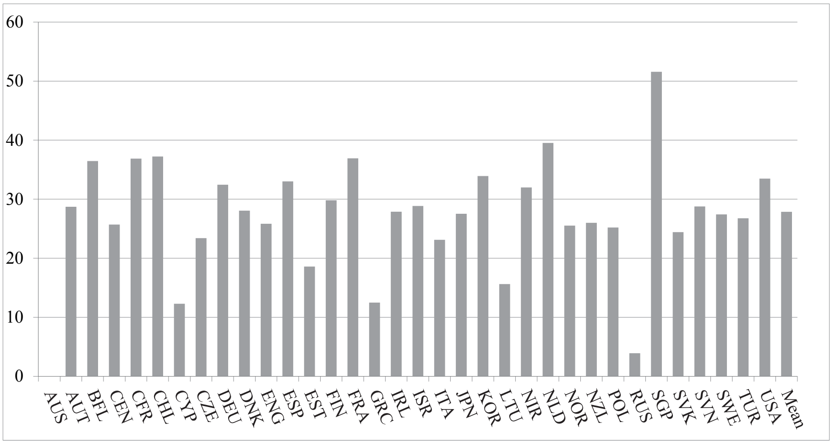

Figure 1 Adjusted R^2s, regression of literacy skills on country specific education variables

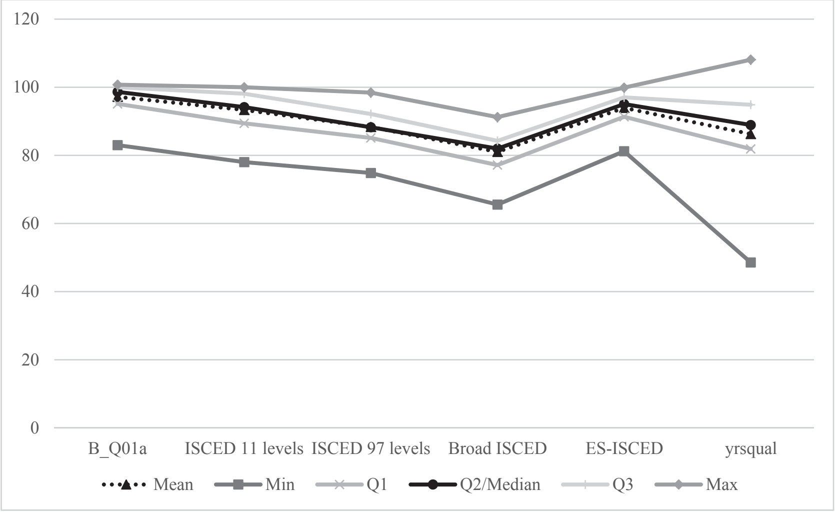

Figure 2 Summary statistics of relative losses in adjusted R^2s predicting literacy skills by comparative education measures

to 100%.[9] Figure 2 shows that the harmonization process from country-specific education variables into the detailed comparable education variable in PIAAC, *B_Q01a*, in itself does not necessarily lead to substantial losses of information and thus explanatory power across countries. Using this variable with up to 14 catego-

9 See Figure 3 and Table 8 in the appendix for absolute rather than relative losses in adjusted R^2. The general picture is the same and conclusions thus apply regardless.

ries, the loss of information is 1.4% on average (median). The next best comparative variable is ES-ISCED (median loss of 4.9%), closely followed by ISCED 2011 levels (5.8%), which however has one category more. All remaining variables lead to median losses of information of more than 10%, with broad ISCED levels performing worst (18%) and ISCED 1997 levels and years of education performing very similarly (11.7 and 11.1% respectively).

However, it is important to also look at the distribution of performance of the different measures across countries, because measures performing very differently across countries are undesirable from a comparability point of view. Using the standard deviation across countries as the summary measure of how differently a comparative education measure captures country-specific information across countries, *B_Q01a* shows the lowest standard deviation of all tested variables (s.d.=3.7, see Table 3). This is followed again by ES-ISCED (s.d.=4.7). ISCED 2011 and 1997 as well as broad levels show higher variation in validity across countries (s.d. of 5.6-6.0). Years of education again come last, with a standard deviation of 12.7.

Especially outliers at the bottom, i.e. countries where a specific measure contains substantially less information than the country-specific education variable, are a matter of concern. Next let's thus look at more detailed results in Table 3 focusing on the strongest losses for B_Q01a, ISCED 2011 levels, and ES-ISCED, i.e. the most promising comparative variables (see shaded cells in Table 3). *B_Q01a* shows the strongest losses for Austria (17%), followed by the Netherlands (9%). With regards to ISCED 2011 levels, the losses are strongest for the Czech Republic (22%), again Austria (18%) and New Zealand (18%). ES-ISCED in contrast produces substantial losses of information for Turkey (19%) and the Czech Republic (18%). These countries are looked at more closely in the following section.

Table 3 Relative adjusted R^2s comparing predictive power of comparative and country-specific education variables predicting literacy skills

Country	k	B_Q01a	ISCED 11 levels	ISCED 97 levels	Broad ISCED	ES-ISCED	yrsqual
AUS*	10	(100)	96.0	90.3	75.4	96.3	85.3
AUT	17	83.0	81.9	78.1	65.6	91.3	71.4
BFL	12	100.0	92.0	87.7	84.4	99.0	86.4
CEN	21	97.2	97.1	85.5	81.1	94.2	89.0
CFR	21	93.1	92.2	82.8	80.1	87.2	83.4
CHL	9	100.0	100.0	95.3	90.7	99.6	98.9
CYP**	14	100.0	100.0	83.7	78.6	99.9	96.8
CZE	13	97.7	78.0	77.2	75.3	81.9	85.4
DEU	16	96.1	95.9	90.4	76.8	96.2	90.1
DNK	14	100.5	94.7	91.6	82.9	96.1	92.9
ENG	29	93.2	88.4	82.0	75.3	90.2	48.6

Country	k	B_Q01a	ISCED 11 levels	ISCED 97 levels	Broad ISCED	ES-ISCED	yrsqual
ESP	12	100.0	99.6	94.4	82.5	96.7	95.7
EST**	19	96.4	94.5	83.0	78.8	94.4	98.2
FIN	12	95.9	95.9	90.7	88.6	97.4	94.5
FRA	17	93.7	87.5	86.1	78.3	90.3	88.9
GRC**	11	100.0	98.5	89.9	84.0	96.7	89.1
IRL	14	99.8	99.8	94.0	86.5	95.3	89.9
ISR	11	99.4	91.5	85.0	82.3	97.6	87.7
ITA	12	100.0	96.8	96.8	81.1	99.6	94.8
JPN	14	100.7	98.1	91.1	90.6	97.1	96.2
KOR	12	99.2	99.2	95.4	91.2	97.9	95.1
LTU**	13	95.5	92.8	87.1	82.9	96.7	85.3
NIR	29	95.1	92.8	87.4	79.8	94.9	52.2
NLD	17	91.3	88.6	86.4	75.6	92.1	81.3
NOR	13	100.0	92.1	89.5	76.1	94.2	81.4
NZL	19	94.8	82.0	74.8	68.6	86.1	76.8
POL	10	99.8	88.7	87.8	85.1	90.7	94.3
RUS**	10	100.0	100.0	95.4	82.3	92.7	60.3
SGP	10	98.1	98.1	93.6	89.3	94.7	95.6
SVK	12	100.0	87.2	86.8	81.7	96.2	79.0
SVN	15	98.6	86.2	84.5	82.6	89.8	94.9
SWE	17	94.9	93.8	92.3	82.4	91.4	86.0
TUR	12	99.9	99.9	98.4	71.9	81.2	80.2
USA***	12	93.6	93.6	88.7	86.3	98.4	108.1
Mean	14.7	97.2	93.3	88.4	81.0	93.9	86.3
Std. deviation	4.8	3.7	5.8	5.6	6.0	4.7	12.7
Min	9	83.0	78.0	74.8	65.6	81.2	48.6
Q1	12	95.1	89.4	85.1	77.2	91.3	81.9
Q2/Median	13	98.6	94.2	88.3	82.0	95.1	88.9
Q3	17	100.0	98.1	92.2	84.3	97.0	94.9
Max	29	100.7	100.0	98.4	91.2	99.9	108.1

Notes. PIAAC rounds 1 and 2 data, complex survey design and plausible values taken into account. k=number of categories in the country-specific education variable. Shaded cells refer to results discussed in more detail in section 4.3.

* Since Australia did not submit country-specific variables to OECD, the predictive power of B_Q01a relative to the country-specific variable cannot be computed for Australia. In the subsequent models, adjusted R^2 relative to the adjusted R^2 of B_Q01a are reported for Australia.

** Countries which have been identified as potentially problematic in the benchmark model (Figure 1).

*** The USA is the only country where years of education explain 8% more variation than the country-specific variable. This is impossible if the yrsqual variable was derived from the country specific variables, as stated in the documentation. Therefore, data processing for this variable must have differed in some way for the USA.

Detailed Country Analyses

For Austria, the loss of information is, with 17%, already quite strong when using in *B_Q01a*. 16 Austrian education categories correspond to 9 *B_Q01a* categories, meaning a substantial amount of aggregation even for the most detailed education variable in PIAAC. Looking at the regression coefficients for the country specific education variable (see Table 4), especially ISCED 3A-B, ISCED 4A-B and ISCED 5B are revealed to be highly heterogeneous comparative education categories in Austria with respect to literacy skills. At ISCED 3A-B, respondents with the lowest qualification, dual system apprenticeship ("Lehre mit Berufsschule"), achieve substantially lower literacy scores (-15 points) than those in the middle category, vocational school ("Fach- oder Handelsschule: 2 Jahre und länger"), and these again substantially lower scores (-23 points) than respondents in the highest and smallest category, general secondary school ("AHS (z.B. Gymnasium)" [10]). The former two are vocational qualifications, the first one involving only part-time schooling, and the second one school-based, and the latter refers to university-preparatory upper secondary education. At ISCED 4A-B, we also find a skill difference of 21 points between the two qualifications classified here, nursing school and vocational college („Berufsbildende Höhere Schule BHS (z.B. HAK, HTL, BAKIP)"). In fact, the literacy skills of nursing school graduates are virtually identical to those of vocational school graduates at ISCED level 3. Given this programme can be entered at age 16, i.e. at a lower age than the usual completion age of ISCED 3A-B, one may wonder whether the qualification is misclassified in ISCED level 4. At ISCED 5B, graduates of the lowest country-specific category, „Meister- und Werkmeisterprüfung, Bauhandwerkerprüfung" (completion of the master crafts exam), achieve the same level of literacy skills as those who completed upper secondary vocational or nursing school, while those with other ISCED 5B qualifications in Austria show 18 to 37 points higher literacy scores (the high scores refer to fairly small categories though). Only the aggregation of the two country specific categories corresponding to ISCED 5A, Master's degree level, does not pose any validity problems since both groups perform rather equally (however, the country-specific variable does not differentiate the type of higher education institution, polytechnic or university, where further heterogeneity may be hidden).

Had other countries differentiated types of education within categories of *B_Q01a* in similar ways, their results in terms of predictive power of *B_Q01a* relative to the country-specific variable might have looked similarly, too: Most country-specific education variables in PIAAC are much less differentiated (see column "k" in Table 3), and the correlation between the number of categories in the national measurement instrument and the loss of information when predicting literacy skill by *B_Q01a* amounts to -.51.

---

10 Acronyms are decoded in Table 4.

The Netherlands is another interesting case to look at, where the most detailed harmonized education variable in PIAAC leads to a loss of 9% of predictive power with regards to literacy skills. Here, also 16 country specific education categories are harmonized into 9 categories. At ISCED level 2 we find 3 country specific categories linked to vastly different average literacy skills (see Table 10 in the appendix). It is in this sense problematic that two tracks in Dutch lower secondary education are classified as 'general education' in ISCED, while one track is actually markedly pre-vocational. Upper secondary education in the Netherlands is also highly stratified, with three qualifications classified as 'ISCED 3C 2 years or more', and another three qualifications classified as ISCED 3A-B. While the lowest category in ISCED 3C shows the same literacy scores as those in pre-vocational ISCED 2, the other two perform substantially higher, but still below those having academic ISCED 2 as their highest attainment. In ISCED 3A-B, the largest and only vocational category performs 21 to 23 points lower than the two smaller general categories. Within tertiary education, which is also tracked in the Netherlands, we again find substantial literacy skill differences within ISCED 5A medium (Bachelor's degree level), between graduates of vocational higher education and traditional universities.[11] It is interesting to note that ISCED 5A, Master's degree level, and ISCED 6 are very close.

For the Czech Republic, the low performance of the ISCED variables that are more aggregated than *B_Q01a* is due to the fact that there are substantial differences in literacy skills between those classified as 'ISCED 3C 2 years or more' and the three categories classified in ISCED 3A-B (see Table 11 in the appendix). Even though vocational, technical and academic ISCED 3A are associated with different literacy skills, their aggregation in B_Q01a does not lead to a substantial loss in predictive power. Summarizing ISCED 2 and '3C shorter than 2 years' in ISCED level 2 in the aggregated ISCED variables (and ES-ISCED) does not pose any problems either. The low performance of ES-ISCED for the Czech Republic lies in the aggregation of upper secondary vocational education, no matter whether it gives access to tertiary education or not. This is the result of the unintended coding of ES-ISCED using orientation rather than destination (see Table 2).

For New Zealand (for detailed results see Table 12 in the appendix), while *B_Q01a* works reasonably well, aggregation to main ISCED 2011 levels again comes at a price. Merging 'ISCED 3C shorter than 2 years' and ISCED 2 leads to a heterogeneous ISCED level 2 in the comparative ISCED variables because those classified as ISCED 2 have on average 25 and 37 points lower literacy scores. However, since these latter individuals do not actually have any educational qualification, while the lowest general school-leaving qualification in NZL is classified as

11 Remember that this differentiation was not made in the Austrian education variable.

Table 4 Detailed regression results for Austria, country-specific variable and B_Q01a

	Austrian educational qualifications				B_Q01a		
	Category (German)	Description in English	b	SE	Category	b	SE
1	Kein Pflichtschulabschluss	No compulsory school	-18.0	9.4	ISCED 1	-21.3	9.5
2	Pflichtschulabschluss	Compulsory school	REF		ISCED 2	REF	
4	Fach- oder Handelsschule: < 2 Jahre	Vocational School (< 2 Years)	14.9	4.6	ISCED 3C <2 years	14.6	4.5
3	Lehre mit Berufsschule	Apprenticeship	13.9	2.3			
5	Fach- oder Handelsschule: 2 Jahre und länger	Vocational School (2 Years and longer)	28.8	2.9	ISCED 3A-B	19.9	2.3
8	AHS (z.B. Gymnasium)	Academic Secondary School	50.3	4.1			
6	Fach- oder Handelsschule: Diplomkrankenpflege	Nursing	27.8	4.6	ISCED 4A-B	44.7	3.2
9	BHS (z.B. HAK, HTL, BAKIP)	Vocational college	48.8	3.4			
7	Meister- oder Werkmeisterprüfung	Master craftsman's certificate	27.2	4.1			
10	Kolleg, Abiturientenlehrgang	Post-secondary courses	64.6	5.9			
11	Akademie (z.B. Pädak, SozAK, BPA, Med.-Tech. Akademie, LW, MilAK)	Post-secondary colleges	45.8	4.1	ISCED 5B	39.6	3.2
12	Universitäre Lehrgänge (ohne vorangegangenes Studium)	University courses	55.4	10.0			

Austrian educational qualifications				B_Q01a		
Category (German)	Description in English	b	SE	Category	b	SE
13 Universität oder Fachhochschule: Bakkalaureat/Bachelor	University-Bachelor	52.4	7.6	ISCED 5A, bachelor degree	52.0	7.5
14 Universität oder Fachhochschule: Magisterium/Master (Diplomstudium, Doktorat als Erstabschluss)	University-Master	61.8	3.3	ISCED 5A, master degree	61.4	3.2
15 Postgraduale Universitätslehrgänge (z.B. MBA, MAS)	Post-graduate courses	58.2	7.1			
16 Doktorat nach akademischem Erstabschluss	Doctoral Programme	54.1	5.4	ISCED 6	54.3	5.4

Notes. Effects of control variables not shown. Acronyms in alphabetical order:
AHS: Allgemein bildende Höhere Schule (general secondary school)
BAKIP: Bildungsanstalt für Kindergartenpädagogik (specialized type of vocational secondary schools, secondary school for nursery-school teaching)
BHS: Berufsbildende Höhere Schule (vocational secondary school, permits university entrance)
BPA: Berufspädagogische Akademie (outdated post-secondary school for vocational school teaching)
HAK: Handelsakademie (specialized type of vocational secondary schools, secondary trade school)
HTL: Höhere Technische Lehranstalt (specialized type of vocational secondary schools, secondary technical school)
LW: probably ‚Landwirtschaftliche Akademie' (agricultural post-secondary school)
MAS: Master of Advanced Studies
MBA: Master of Business Administration
MilAK: Militärakademie (military post-secondary school)
PädAK: Pädagogische Akademie (outdated post-secondary school for nursery-school and primary school teaching)
SozAK: Akademie für Sozialarbeit (outdated post-secondary school for social work)

'ISCED 3C short',[12] one may also wonder whether the ISCED mapping for NZL is comparable with that of most other countries, where the first school-leaving qualification is awarded at the end of ISCED level 2 and not having any qualification is regarded as ISCED 1 if the number of years of schooling required for completion of ISCED 1 is fulfilled (otherwise ISCED 0). Furthermore, at ISCED level 3, qualifications classified as 'ISCED 3C 2 years or more' are related to substantially lower literacy skills than qualifications classified as ISCED 3A-B (differences of up to 40 points).

Turkey shows up to be problematic in two of the categorical comparative variables only, namely broad ISCED levels and ES-ISCED. Why is this so? Both variables drop the distinction between ISCED levels 0 and 1, which is still very relevant in less developed countries. Given the lower level of educational attainment of the Turkish population (see Table 5 in the appendix), and the consequently rather important distinction between ISCED levels 0 and 1 also in terms of literacy skills, it would thus be better for ES-ISCED to not drop the distinction between ISCED 0 and 1 whenever including less developed countries in empirical analyses of education effects.

To summarize, while ISCED 2011 works better in many countries than ISCED 1997, aggregating 'ISCED 3C 2 years or more' with ISCED 3 A-B remains a problematic aggregation (see example for the Czech Republic and New Zealand here). Countries like Austria, where apprenticeship training gives access to tertiary education, show similar problems *within* ISCED 3 A-B. Upper secondary education in developed countries is too heterogeneous in terms of skill production due to content, quality and place of learning to be meaningfully represented by one single educational attainment category. Access to tertiary education (including short cycle and even master crafts programs) may not be the best criterion to render categories comparable across countries. The tertiary qualification that allows the classification of apprenticeships as ISCED 3B in Austria, the master crafts certificate, actually also does not fit in in terms of skills, so this coding may actually be the underlying culprit. Countries with tracked school systems like the Netherlands would benefit from a more differentiated ISCED level 2, and for countries with low or late educational expansion like Turkey, the distinction between ISCED 0 and 1 remains important. Finally, the completion of various school grades without qualification is classified differently across countries (see the example of New Zealand), leading to comparability problems at the lower end of the ISCED classification.

12 Educational programmes with destination C usually only prepare for the labour market. The classification of the first general school leaving certificate in New Zealand as 'ISCED 3C shorter than 2 years' strongly reminds of the disputable classification of the respective UK qualifications (see Schneider, 2008).

Conclusions and Recommendations

Respondent's educational attainment is probably the most important single variable in the PIAAC background questionnaire, used as a predictor of adult skills, labor market outcomes, and control variable. This study evaluated a range of comparative education measures, mostly based on ISCED, with respect to their predictive validity when using skills as validation variable, which has not been done before.

At a theoretical level, the way that ISCED is implemented in cross-national surveys, including PIAAC, often does not allow studying the antecedents and consequences of educational attainment with respect to program content (orientation), quality (destination), or place of learning, even though these are important elements when studying skill acquisition and labor market outcomes. Furthermore, skill selectivity, academic demand or place of learning that is not expressed in program orientation or destination as defined in ISCED can be shown to be important within countries (see the results for Austria regarding apprenticeship and school-based vocational education, and lower secondary school tracking in the Netherlands) but are not represented in any version of ISCED. For future cycles of PIAAC, and surveys where education is used as an indicator for general basic skills, it is thus important that general and vocational educational qualifications can be clearly distinguished and classified, and that further dimensions of education are reflected, such as place of learning and quality in terms of selectivity.

Empirically, with some exceptions, the most detailed comparative education variable in PIAAC, *B_Q01a*, works rather well as a harmonized education measure. It well reflects quantity and partially also quality of education, but disregards content (vocational vs. general) unless this overlaps with quality. Aggregating to ISCED levels (2011 and especially 1997) leads to substantial reductions of comparative construct validity and thus comparability, which illustrates that quantity of education is an important dimension, but not sufficient. The implementation of ISCED in *B_Q01a* is thus definitely an advantage compared to using ISCED 1997 main levels only, as is e.g. done in the European Union Survey of Income and Living Conditions (EU-SILC) and recommended in the Core Social Variables (European Commission 2007). The validation analyses also show that ISCED 2011 main levels are substantially better suited for the multivariate analysis of adult skills than ISCED 1997 main levels, owing to the better reflection of quantity and content at the tertiary level. 'Broad' ISCED levels (low, medium, high) do not even reflect the quantity of education sufficiently. The analyses also show that if you aggregate detailed education categories in a way that keeps the important dimension of content (vocational vs. general) and drops less important distinctions regarding quantity, like in ES-ISCED, one can achieve acceptable harmonization results with a variable containing just eight categories. Years of education in contrast do not well represent the skill information contained in country-specific education categories, and they also do so quite differently across countries. Reducing educational attainment to its quantity dimension is thus not recommendable when trying to proxy skills (however, the relationship between education and literacy skills is, on average, moderate rather than strong, and thus ISCED not a good proxy for skills anyway, see Massing & Schneider, 2017).

A limitation of this study, especially concerning the rather positive result for *B_Q01a*, lies in the already mixed quality of the country-specific measures in PIAAC. They are often no more detailed than *B_Q01a* – many country teams have implemented questionnaire items that just minimally satisfy the requirements of the comparative PIAAC variable *B_Q01a*, rather than measuring education at the level of detail that would have been most suitable for the respective national education system. If more countries measured educational attainment in more detail, the results would potentially look a little less positive for *B_Q01a*. Indeed, when limiting the results reported in section 4.2 to countries that have at least two country-specific categories merged into one category of B_Q01a[13] - a very minimal and conservative indicator of quality - the average loss of information of B_Q01a amounts to 4.1% on average (compared to 2.8% when including all countries).

A further limitation of the study may be the inclusion of sex and age as control variables in all models, in combination with relative R^2s as the indicator for comparative validity: If countries differ in the partial R^2 of age and gender, comparative validity (the relative reduction in R^2 due to education harmonization) will be biased, and will be biased more the higher the partial R^2 of age and gender. The effect of gender on skills is however generally low, and the effect of age is to a substantial degree due to educational attainment (OECD, 2016a). With this in mind, and given the consistency between relative and absolute losses in R^2s, and the fact that this bias is a downward (i.e. conservative) bias, it is very unlikely that the exclusion of controls from all models would substantially change the conclusions: if anything, they would become stronger.

For secondary data analyses of PIAAC and other cross-national survey data involving educational attainment, it can be concluded that in order to avoid confounding, improve validity and thereby also comparability, education is best measured using a coding scheme that is neither too differentiated to make the analyses overly cumbersome, nor too simplified. ES-ISCED or ISCED 2011 levels can both be used, and theoretical considerations should be used in the decision for one or the other. Further aggregations should always be accompanied by sensitivity checks, comparing statistical results when using more and less detailed education variables, in order to make sure that the results of comparative survey research are valid and

13 These countries are AUT, CEN, CFR, CHL, CZE, DEU, ENG, EST, FIN, FRA, IRL, ISR, KOR, LTU, NIR, NLD, NZL, POL, SGP, SVN, SWE, TUR, USA. They have 16 education categories on average, while the remaining countries (BFL, CYP, ESP, ITA, JPN, NOR, RUS, SVK, DNK, GRC) have 12.

not due to measurement and harmonization artefacts. Ideally, ISCED would be implemented in a better way in comparative surveys, paying more attention to the dimensions of education to be measured. Even more ideally, ISCED itself would be revised again in the near future so as to better reflect the various dimensions of education.

References

Allmendinger, J. (1989). Educational systems and labor market outcomes. *European Sociological Review*, *5*(3), 231–250.

Avvisati, F., & Keslair, F. (2017). REPEST: Stata module to run estimations with weighted replicate samples and plausible values. https://EconPapers.repec.org/RePEc:boc:bocode:s457918.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education* (first). Chicago (IL), London: University of Chicago Press.

Blau, P. M., & Duncan, O. D. (1967). *The American occupational structure*. New York, London: Wiley.

Braun, M., & Mohler, P. P. (2003). Background variables. In J. A. Harkness, F. J. R. van de Vijver, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 101–116). Hoboken (NJ): J. Wiley.

Braun, M., & Müller, W. (1997). Measurement of education in comparative research. *Comparative Social Research*, *16*, 163–201.

Breen, R., Luijkx, R., Müller, W., & Pollak, R. (2009). Nonpersistent inequality in educational attainment: Evidence from eight European countries. *American Journal of Sociology*, *114*(5), 1475–1521.

Breen, R., Luijkx, R., Müller, W., & Pollak, R. (2010). Long-term trends in educational inequality in europe: Class inequalities and gender differences. *European Sociological Review*, *26*(1), 31–48. https://doi.org/10.1093/esr/jcp001

Bukodi, E., Róbert, P., Szilvia, A., & Altorjai, S. (2008). The Hungarian educational system and the implementation of the ISCED-97. In S. L. Schneider (Ed.), *The International Standard Classification of Education (ISCED-97). An evaluation of content and criterion validity for 15 European countries* (pp. 200–215). Mannheim: MZES.

Ehling, M. (2003). Harmonising data in official statistics. In J. H. P. Hoffmeyer-Zlotnik & C. Wolf (Eds.), *Advances in cross-national comparison: A European working book for demographic and socio-economic variables* (pp. 17–31). New York; London: Kluwer Academic/Plenum.

European Commission. (2007). *Task force on core social variables final report*. Luxembourg: European Commission.

Granda, P., Wolf, C., & Hadorn, R. (2010). Harmonizing Survey Data. In J. A. Harkness, M. Braun, B. Edwards, T. P. Johnson, P. P. Mohler, B.-E. Pennel, & T. W. Smith (Eds.), *Survey methods in Multinational, Multiregional, and Multicultural Contexts* (pp. 315–332). Hoboken (NJ): Wiley.

Hall, P., & Soskice, D. (2001). An Introduction to Varieties of Capitalism. In P. Hall & D. Soskice (Eds.), *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage* (pp. 1–56). Oxford: Oxford University Press. https://doi.org/10.1093/0199247757.001.0001

Haller, M., König, W., Krause, P., & Kurz, K. (1985). Patterns of Career Mobility and Structural Positions in Advanced Capitalist Societies: A Comparison of Men in Austria, France, and the United States. *American Sociological Review*, *50*(5), 579–603. https://doi.org/10.2307/2095376

Heisig, J. P., & Solga, H. (2015). Secondary Education Systems and the General Skills of Less- and Intermediate-educated Adults: A Comparison of 18 Countries. *Sociology of Education*. https://doi.org/10.1177/0038040715588603

Hoffmeyer-Zlotnik, J. H. P., & Wolf, C. (Eds.). (2003). *Advances in cross-national comparison: A European working book for demographic and socio-economic variables*. New York; London: Kluwer Academic/Plenum.

Kerckhoff, A. C., & Dylan, M. (1999). Problems with international measures of education. *Journal of Socio-Economics*, *28*(6), 759–775.

Kerckhoff, A. C., Ezell, E. D., & Brown, J. S. (2002). Toward an improved measure of educational attainment in social stratification research. *Social Science Research*, *31*(1), 99–123.

Kieffer, A. (2010). Measuring and Comparing Levels of Education: Methodological Problems in the Classification of Educational Levels in the European Social Surveys and the French Labor Force Surveys. *Bulletin de Méthodologie Sociologique*, *107*(1), 49–73. https://doi.org/10.1177/0759106310369974

König, W., Lüttinger, P., & Müller, W. (1988). *A comparative analysis of the development and structure of educational systems*. Mannheim.

Massing, N., & Schneider, S. L. (2017). Degrees of Competency: The Relationship between Educational Qualifications and Adult Skills across Countries. *Large-Scale Assessments in Education*, *5*(1), 1–34. https://doi.org/10.1186/s40536-017-0041-y

Müller, W., & Karle, W. (1993). Social selection in educational systems in Europe. *European Sociological Review*, *9*(1), 1–23.

Müller, W., & Klein, M. (2008). Schein oder Sein: Bildungsdisparitäten in der Europäischen Statistik. Eine Illustration am Beispiel Deutschlands. *Schmollers Jahrbuch*, *128*(4), 511–543.

OECD. (2013). *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*. Paris: OECD Publishing. https://doi.org/10.1787/9789264204256-en

OECD. (2016a). *Skills Matter: further results from the Survey of Adult Skills*. Paris: OECD. https://doi.org/10.1787/9789264258051-en

OECD. (2016b). *Technical Report of the Survey of Adult Skills (PIAAC)* (2nd ed.). Paris: OECD.

Przeworski, A., & Teune, H. (1970). *The logic of comparative social inquiry. The logic of comparative social inquiry* (Vol. 1). New York: Wiley-Interscience.

Saar, E. (2008). The Estonian educational system and the ISCED-97. In S. L. Schneider (Ed.), *The International Standard Classification of Education (ISCED-97). An evaluation of content and criterion validity for 15 European countries* (pp. 237–252). Mannheim: MZES.

Schneider, S. L. (2008). The application of the ISCED-97 to the UK's educational qualifications. In S. L. Schneider (Ed.), *The International Standard Classification of Education*

(ISCED-97). An evaluation of content and criterion validity for 15 European countries (pp. 281–300). Mannheim: MZES. https://doi.org/10.13140/RG.2.1.1993.5127

Schneider, S. L. (2010). Nominal comparability is not enough: (In-)equivalence of construct validity of cross-national measures of educational attainment in the European Social Survey. *Research in Social Stratification and Mobility*, *28*(3), 343–357. https://doi.org/10.1016/j.rssm.2010.03.001

Schneider, S. L. (2013). The International Standard Classification of Education 2011. In G. E. Birkelund (Ed.), *Class and Stratification Analysis. Comparative Social Research.* (Vol. 30, pp. 365–379). Bingley: Emerald. https://doi.org/10.1108/S0195-6310(2013)0000030017

Schneider, S. L. (2016). *The Conceptualisation, Measurement, and Coding of Education in German and Cross-National Surveys* (GESIS Survey Guidelines). *GESIS Survey Guidelines.* Mannheim: GESIS – Leibniz Institute for the Social Sciences. https://doi.org/doi: 10.15465/gesis-sg_en_020

Schneider, S. L., Joye, D., Wolf, C., & Surveys, C. (2016). When Translation is not Enough: Background Variables in Comparative Surveys. In C. Wolf, D. Joye, T. W. Smith, & Y.-C. Fu (Eds.), *The SAGE Handbook of Survey Methodology* (pp. 288–307). Los Angeles.

Smith, T. W. (1995). Some aspects of measuring education. *Social Science Research*, *24*(3), 215–242. https://doi.org/10.1006/ssre.1995.1008

Smith, T. W. (2011). Refining the Total Survey Error Perspective. *International Journal of Public Opinion Research*, *23*(4), 464–484. https://doi.org/10.1093/ijpor/edq052

Straková, J. (2008). The Czech educational system and evaluation of the ISCED-97 implementation. In S. L. Schneider (Ed.), *The International Standard Classification of Education (ISCED-97). An evaluation of content and criterion validity for 15 European countries* (pp. 216–225). Mannheim: MZES.

UNESCO Institute for Statistics. (2012). *International Standard Classification of Education - ISCED 2011.* Montreal: UNESCO Institute for Statistics.

Weber, M. (1922). *Wirtschaft und Gesellschaft*. Tübingen: J.C.B Mohr (Paul Siebeck).

Wolf, C., Schneider, S. L., Behr, D., & Joye, D. (2016). Harmonizing survey questions between cultures and over time. In C. Wolf, D. Joye, T. W. Smith, & Y. Fu (Eds.), *The SAGE Handbook of Survey Methodology* (pp. 502–524). Los Angeles: Sage.

Online Appendix

http://mda.gesis.org/index.php/mda/article/view/2017.15